Status and issues of the protoDUNE DQM and interfaces

Maxim Potekhin

DUNE Weekly @ BNL

04/12/2017

Overview

- Mission statement update for DQM (next slide)
  - Beam Instrumentation is a new factor in DQM
- See the “p3s plan” attached to this agenda (requested by Flavio and Christos)
- Beam Instrumentation group meetings are now on Indico and Zoom
  - https://indico.fnal.gov/categoryDisplay.py?categId=587
- Progress with p3s, both with functionality and deployment
- CERN Twiki for protoDUNE computing has been updated and reformatted (Maxim)
  - https://twiki.cern.ch/twiki/bin/view/CENF/DUNEProtSPComputing
- Also bookmark the “central” CENF computing page:
  - https://twiki.cern.ch/twiki/bin/view/CENF/Computing
- Other news: Ruth Pordes (FNAL) has started in her role of the protoDUNE computing liaison at CERN
  - Ruth was the director of the Open Science Grid Consortium for many years
  - proposed protoDUNE data challenges in October 2017 and early 2018, scope and schedule TBD
  - will be at CERN the week of the 17th for a series of meetings (including with myself)
  - will be stationed at CERN continuously for ~1 year

DQM mission update

- “prompt processing” has been renamed “DQM” (according to Ruth)
- Relax the nominal 10 min latency requirement up to an hour in order to allow more complex payload to run in p3s
- Leaning towards EOS for input and output data (will impact p3s placement)
- Add a requirement for the Beam Instrumentation (BI) data inclusion in the DQM process
- New function of DQM is to validate the trigger and BI in general
  - new software in addition to what was planned before
- The BI data gets recorded into the LHC accelerator data DB system and will have to be merged with the TPC/PD data by matching timestamps
  - see next slide

Beam Instrumentation

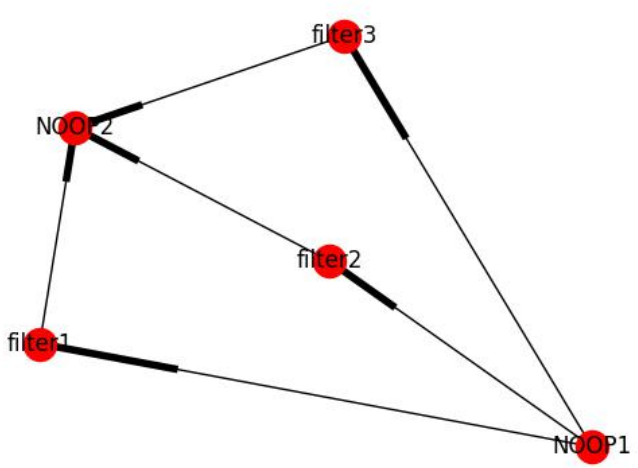
- Beam Instrumentation has a few components and at least two different routes for data:
  - **Jon Paley**: *'The fiber tracker data, other spill-relevant/monitoring data, and ckov data will all go into the BI DB. The pLAPPD ToF system (from FNAL) will be fully integrated with the TPC DAQ, so the data will go directly into the artdaq stream. The other ToF system (from CERN) is "up in the air".'*
- The instrumentation database:
  - Common for all of CERN accelerator division, high performance,complex setup with specific access modes (to guaranteee performance and availability), large data rate and volume
  - Back end: Oracle RAC (but RDBMS not exposed)
  - APIs: Java-based - will need to be bridged to PostgreSQL or some other medium
  - The BI data will be available at the end of spill by querying the DB
  - essentially time-series data for the fiber tracker, Cherenkov and potentially beam monitors
  - DB Schemas have not been created, there was a mention of late 2017 - imho way too late - need specifications/prototypes/mockups soon
- Beam Instrumentation data merge is based on timestamps i.e. is time series data:
  - exact algorithm and data handling for this step is unclear
  - may potentially be a bottleneck
- This is responsibility of the Beam Instrumentation Group and outside of the purview of Prompt Processing, interfaces and algorithms are part of offline
  - yet a dependency for DQM

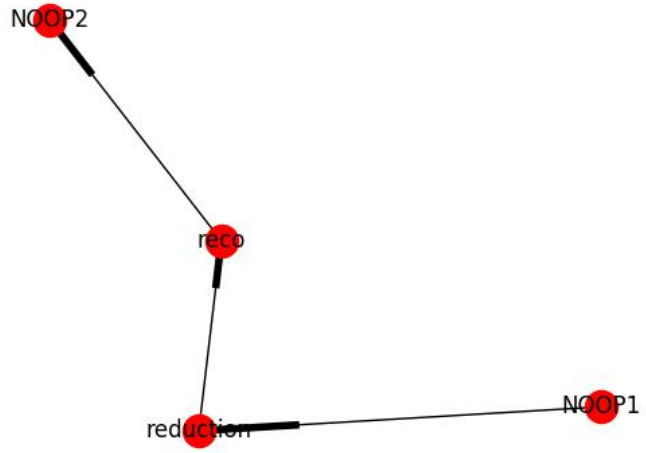
P3S status and plans

- The system has been continuously improved over the past weeks
  - UI and functionality improvements
  - better error handling
  - resolution of concurrency and deployment issues
- Deployed on the “neutdqm” test cluster at CERN on Apache/PostgreSQL - 10 machines total with one hosting Apache and DB, the rest of the nodes acting as worker nodes
  - continuous “burn-in” testing with trivial test payloads, system consistency checks
  - pilot submission via pdsh
- Currently access from BNL to the system Web Interface is only possible by ssh tunneling via lxplus
  - DAQ team took it upon themselves to sort out the issue of proxies/ports etc, but no progress yet
- Will need to follow up with CERN team on explanding to the full core count of ~300
- Shift to realistic (i.e. LArSoft-based) payloads
  - very useful to testing, useful as a way to manage current MC production (Robert and Dorota)
- EOS interface to be tested
- lxbatch deployment (i.e. Tier-0) pending installation of AutoPyFactory and other components (separate effort needed)

P3S workflow support

- p3s supports both individual jobs and workflows as first-class objects
  - workflow execution is automatic, with jobs triggered when conditions are satisfied
- Workflows are modeled as graphs (DAGs), there is a template library on p3s
- The system uses an industry standard XML schema (GraphML) so a user can create his/her workflow templates in any text editor
  - plus, a number of third party graphic tools is also available to manipulate and visualize graphs
  - a workflow template can be patched by the user in order to reuse same topology but with different payloads and parameters





p3s screenshots (landing page/dashboard)

p3s

DAGs Workflows Jobs Pilots Data Data Types

Summary		System	
Object	Number	Attribute	Value
Pilots: total/idle/running/stopped	19/0/1/18	Current time	04/10/17 15:44:26 America/New_York
Jobs: total/defined/running/finished	14/1/1/12	Server	serenity
Workflows	1	Uptime	4 days, 5:49:46.750000
Datasets	2	>	>

p3s screenshots (pilots running at CERN)

pilots (total in DB: 11)

Pilot States (select one):

- [x] All
- [ ] Active
- [ ] Running
- [ ] Stopped
- [ ] Timed out
- [ ] No Jobs

Submit

ID	Jpid	State	Status	Site	Host	created	registered	last heartbeat	jobs	Jobs Done	Pid
9	15952	running	OK	default	neutdqm09.cern.ch	20170411 19:50:16	20170411 19:50:16	20170411 19:55:58	0	—	15787
7	—	no jobs	OK	default	neutdqm08.cern.ch	20170411 19:50:16	20170411 19:50:16	20170411 19:56:18	0	—	13944
10	—	no jobs	OK	default	neutdqm02.cern.ch	20170411 19:50:16	20170411 19:50:16	20170411 19:56:18	0	—	3816
13	—	no jobs	OK	default	neutdqm04.cern.ch	20170411 19:50:16	20170411 19:50:16	20170411 19:56:18	0	—	12975
12	—	no jobs	OK	default	neutdqm07.cern.ch	20170411 19:50:16	20170411 19:50:16	20170411 19:56:18	0	—	11591
11	—	no jobs	OK	default	neutdqm03.cern.ch	20170411 19:50:16	20170411 19:50:16	20170411 19:56:18	0	—	3224
14	—	no jobs	OK	default	neutdqm01.cern.ch	20170411 19:50:16	20170411 19:50:16	20170411 19:56:18	0	—	13861
4	—	nonstarter	FAIL	default	neutdqm01.cern.ch	20170411 18:32:12	20170411 18:32:12	20170411 18:32:12	0	—	11362
5	11401	stopped	OK	default	neutdqm01.cern.ch	20170411 18:35:42	20170411 18:35:42	20170411 18:37:24	1	308007ca-1f07-11e7-a220-001d0967b087	11399
6	4051	running	OK	default	neutdqm05.cern.ch	20170411 19:50:16	20170411 19:50:16	20170411 19:55:57	0	—	3897
8	3450	running	OK	default	neutdqm06.cern.ch	20170411 19:50:16	20170411 19:50:16	20170411 19:55:57	0	—	3297

Current time: 04/11/17 19:56:24 EST

p3s screenshots (status of jobs on the neutdqm cluster)

jobs (total in DB: 5)

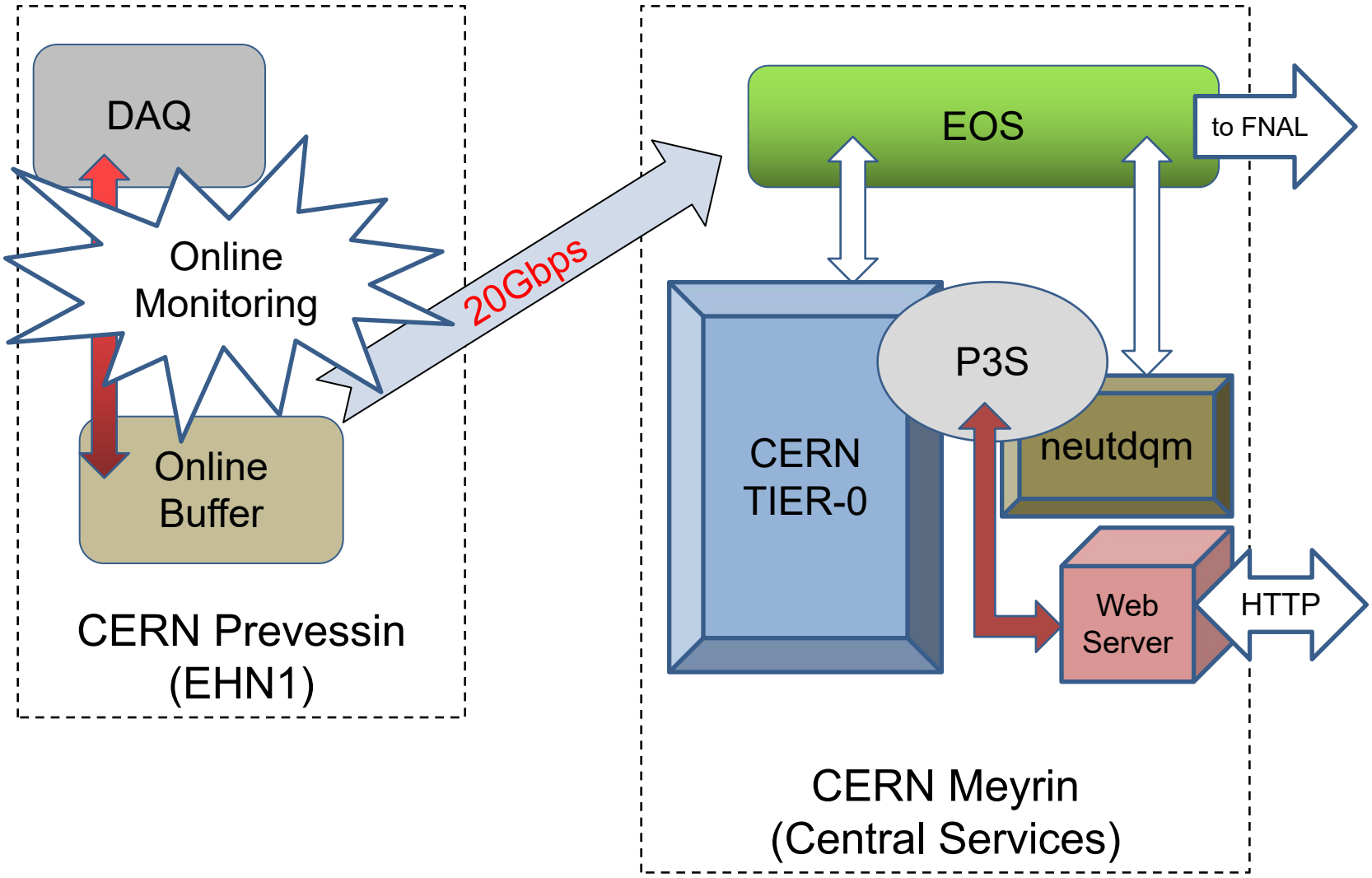
Job States (select one):

- [x] All
- [ ] Template
- [ ] Defined
- [ ] Running
- [ ] Finished
- [ ] Pilot Timed Out

Submit

ID	Uuid	User	Name	Wfuuid	type	Payload	Params	Pri.	t.limit	State	defined	started	stopped	Pid	Errco
7	644ee502-1f12-11e7-9f74-001d096774dd	mxp	sleep1000	—	sleep	/mnt/nas01/users /mxp/projects /p3s/inputs /jobs/sleep100.sh	—	1	100	running	20170411 19:55:51	20170411 19:55:57	—	15952	—
3	b3378608-1f06-11e7-b642-001d0967b087	mxp	sleep1000	—	sleep	/home/maxim /projects /p3s/inputs /jobs/sleep100.sh	—	1	100	nonstarter	20170411 18:32:09	20170411 18:32:12	—	—	—
4	308007ca-1f07-11e7-a220-001d0967b087	mxp	sleep1000	—	sleep	/mnt/nas01/users /mxp/projects /p3s/inputs /jobs/sleep100.sh	—	1	100	finished	20170411 18:35:39	20170411 18:35:43	20170411 18:37:24	11401	0
5	62636506-1f12-11e7-9fb8-001d096774dd	mxp	sleep1000	—	sleep	/mnt/nas01/users /mxp/projects /p3s/inputs /jobs/sleep100.sh	—	1	100	running	20170411 19:55:47	20170411 19:55:57	—	4051	—
6	636aa040-1f12-11e7-9da5-001d096774dd	mxp	sleep1000	—	sleep	/mnt/nas01/users /mxp/projects /p3s/inputs /jobs/sleep100.sh	—	1	100	running	20170411 19:55:49	20170411 19:55:57	—	3450	—

p3s access to NP04 data and location of resources



Misc issues and summary

- The level of collaboration with the Online Monitoring group is lower than has been hoped for, so far OM and DQM appear to be on different tracks
- Reusing and sharing infrastructure components between OM and DQM will depend on decisions on systems architecture to be made in the following weeks
  - meetings at CERN next week
- Scarcity of hardware, with DQM moslty relying on “neutdqm” machine as servers
- p3s is run almost continuously at CERN on the neutdqm cluster, can also be deployed on the “main” neutdqm, will switch to real payloads in the coming weeks
- deployment on Tier-0 will require more work
- As planned p3s will be in place (in some capacity) for the DAQ vertical slice testing this summer, its function to be clarified with the DAQ group
- There is interest in DRA group to utilize p3s as workflow management tool for MC etc
- Interface to the BI data and reconstruction software for its use need to be managed
  - Ruth, Jon